Video Interactions in Online Video Social Networks

FABRÍCIO BENEVENUTO, TIAGO RODRIGUES, VIRGILIO ALMEIDA, and JUSSARA ALMEIDA
Federal University of Minas Gerais, Brazil
and
KEITH ROSS
Polytechnic Institute of NYU

This article characterizes video-based interactions that emerge from YouTube's video response feature, which allows users to discuss themes and to provide reviews for products or places using much richer media than text. Based on crawled data covering a representative subset of videos and users, we present a characterization from two perspectives: the video response view and the interaction network view. In addition to providing valuable statistical models for various characteristics, our study uncovers typical user behavioral patterns in video-based environments and shows evidence of opportunistic behavior.

Categories and Subject Descriptors: H.3.5 [**Information Storage and Retrieval**]: Online Information Services—*Web-based services*; J.4 [**Social and Behavioral Sciences**]

General Terms: Human Factors, Measurement

Additional Key Words and Phrases: Video interactions, video communication, social networks, social media, opportunistic behavior, YouTube, video spam, promotion

ACM Reference Format:

Benevenuto, F., Rodrigues, R., Almeida, V., Almeida, J., and Ross, K. 2009. Video interactions in online video social networks. ACM Trans. Multimedia Comput. Commun. Appl. 5, 4, Article 30 (October 2009), 25 pages.
DOI = 10.1145/1596990.1596994 http://doi.acm.org/10.1145/1596990.1596994

1. INTRODUCTION

Video content is becoming a predominant part of user's daily lives on the Web. By allowing users to generate and distribute their own content to large audiences, the Web has been transformed into a major channel for the delivery of multimedia, leading society to a new multimedia age. Video pervades the Internet and supports new types of interaction among users, including video forums, video chats, video mail, and video blogs. Additionally, a number of services in the current Web 2.0 are offering video-based functions as alternative to text-based ones, such as video reviews for products (www.amazon.com and www.expotv.com are examples of sites that allow users to post video reviews about products), video ads and video responses [Shannon 2007]. This huge growth of multimedia content in the Web is mostly

This research is partially funded by the Brazilian National Institute of Science and Technology for Web Research (MCT/CNPq/INCT Web Grant Number 573871/2008-6) and by research project REBU (MCT/CNPq/CTInfo Grant Number 550995/2007-2).
Authors' addresses: F. Benevenuto, email: fabricio@dcc.ufmg.br; T. Rodrigues, email: tiagorm@dcc.ufmg.br; V. Almeida, email: virgilio@dcc.ufmg.br; J. Almeida, email: jussara@dcc.ufmg.br; K. Ross, email: ross@poly.edu.

due to the evolution of the user from content consumer to content creator. As a consequence, several multimedia issues need to be revisited.

In fact, a recent discussion on the needs and challenges of multimedia research in the context of Web 2.0 pointed out that understanding how users typically behave (e.g., which interactions they establish) is of great relevance [Boll 2007]. As an example, the design of effective video content classification mechanisms is crucial for automatic identification of videos with malicious content such as infringing copyright, pornography, and spam. However, content classification based solely on the raw content can be a challenging research problem due to the typically low quality of user generated videos [Boll 2007] and the multitude of strategies one can make use of to publicize malicious content in a video (URL to website, static image, or streaming). In contrast, understanding how users interact with each other in an online video social network service may highlight aspects inherent to how malicious users behave, which in turn, may be used in a much more effective way for detecting and possibly removing malicious or unwanted content [Benevenuto et al. 2008b, 2009a].

In this article, we perform a large and representative characterization of users interacting with each other essentially via video objects. Particularly, we characterize the use of YouTube's video response feature, which allows one user to video respond to another user's video contribution, creating asynchronous multimedia conversations. The video response feature became a new trend in online video social network systems as a means to exchange knowledge and express ideas through video interactions. Video responses allow users to provide reviews for products or places, and to exchange their opinions about certain themes using a much richer media than simple text.

The characterization of video interactions through video responses is of interest for two reasons. The first is technical, stemming from the necessity to understand video communication in order to evaluate new design choices for video services. The second is sociological, relating to social networking issues that influence the behavior of users interacting primarily via streaming objects, instead of textual content traditionally available on the Web. The video media type opens new doors for originality and spontaneity, but also for content pollution (e.g., spam, promotion, etc.), in the interactions among users of an online social network.

Our characterization of video interactions relies on a large sample collected from YouTube, the currently most popular online video social network. We analyze the characteristics of video responses as well as of the interactions triggered by the use of this feature. We further investigate the presence of oportunistic activities. In summary, the main contributions of this work are as follows:

—*Characterization of the usage of the video response feature*. We characterize the usage of the video response feature, unveiling interesting user behavior and providing statistical models for various characteristics (e.g., popularity profiles and video duration). Such models and findings provide valuable insights for the future design of realistic synthetic workloads for online video social networks.

—*Characterization of video interactions*. In order to understand the characteristics of video-based interactions, we adopt a hierarchical approach to analyze the collected data, looking at the results of the crawls as a two-level hierarchy. First, we focus on the lower level, that is on user interactions engendered by a single responded video. Then we explore the top-level interactions, namely the social network created by the interactions among users across all responded videos. Characterizing interaction in both levels is valuable in order to build models that describe video-based communication. Not only is it important to identify the intrinsic properties of this type of communication in social networking environments, but it is also important to explain how and why these properties emerge. Such an understanding would allow us to build Web systems and services that are both more efficient and user friendly.

—*Evidence of opportunistic behavior in online video social networks*. Our study unveils evidence of two types of opportunistic behavior by YouTube users, namely promotion and video spamming. The former consists of artificially boosting a video rank to make it appear in the top lists provided by YouTube, thus gaining visibility to the content or to its contributor. The latter consists of adding an unrelated video as response to another video in order to promote specific content, advertise to generate sales or disseminate pornography. Additionally, we propose the use of the PageRank algorithm [Brin and Page 1998] in our video response user graph to detect users trying to promote a video by boosting its position in lists of most responded videos provided by YouTube.

This article is organized as follows. The next section discusses related work. Section 3 describes how we crawled YouTube to create a graph of video interactions among users. Section 4 presents statistics of video responses, whereas the following two sections present the characterization of video interactions at the two levels of granularity, namely the video interaction level and the network interaction level. Section 7 discusses opportunistic behavior in video interactions. Finally, Section 8 concludes the article and presents possible directions for future work.

2. RELATED WORK

Over the last few years, there have been a number of studies that explored the various aspects of social networking sites. For example, the efforts in Cha et al. [2007], Mislove et al. [2007], Gill et al. [2007], Golder et al. [2007], and Chua et al. [2009] explored the overall scope, structure, and friend relationship patterns of popular online social networks such as Flickr,[1] YouTube, LiveJournal,[2] Facebook,[3] and Orkut.[4] Particularly, an interesting study of YouTube is presented in [Cha et al. 2007]. The authors analyzed the popularity distribution, popularity evolution, and content characteristics of YouTube and of a Korean video sharing service. They also analyzed system issues that could be used to improve video distribution mechanisms, such as caching and peer-to-peer distribution schemes.

Additionally, the authors in Gill et al. [2007] present a characterization of the YouTube traffic collected from a university campus network, comparing its properties with those previously reported for Web and media streaming workloads. They found that HTTP GET requests, used for fetching content from the server, correspond to over 99% of the total requests sent to the server, and that requests sent from the campus to the server follow typical daily and weekly patterns. They also analyzed file sizes, video durations, video bit rates, video ages, video ratings, and video categories, comparing these properties with those of objects in other media types retrieved from YouTube as well as of traditional Web and media streaming workloads. Another characterization of the YouTube traffic collected from a university campus is presented in Zink et al. [2008]. Based on their measurements, the authors designed trace-driven simulations to show that client-based local caching, P2P-based distribution, and proxy caching can significantly reduce network traffic and allow faster access to videos. These studies show evidence of significant differences in user and video access patterns compared with traditional Web servers.

In terms of textual interactions, Krishnamurthy et al. [2008] analyzed an online social network formed by users on Twitter,[5] a short message service (SMS). Twitter messages can be received on the cell phone as e-mail, RSS feed, and so forth. The authors examined geographical distribution and user behavior in this environment. More recently, Huberman et al. [2009] showed that Twitter users

[1]www.flickr.com.
[2]www.livejournal.com.
[3]www.facebook.com.
[4]www.orkut.com.
[5]www.twitter.com.

have a very small number of friends compared to the number of followers and followees they declare. Leskovec and Horvitz [2008] analyzed the network formed by instant messages in Microsoft Messenger, finding that it is highly connected, with 99.9% of the nodes belonging to the largest weakest connected component, and exhibits small-world properties. They also found a strong correlation between instant message communication and user geographic location. A network formed by textual communication in Cyworld,[6] a popular Korean online social network, was recently approached in Chun et al. [2008]. The authors compared the explicit friend relationship network with the implicit network created by messages exchanged on Cyworld's guestbook, finding several similarities in terms of network structure. Particularly, they found that the in-degree and the out-degree distributions are close to each other, and that social interactions through the guestbook are very well reciprocated. Finally, a framework to measure user interest in conversations about a certain YouTube video is proposed in Choudhury et al. [2009].

A network formed by mobile communication was studied in Onnela et al. [2007]. The authors examined communication patterns of people on mobile phones, arguing that the stability of the communication network largely depends on the weak ties in the network. Additionally, a social network formed by phone calls was studied in Seshadri et al. [2008]. Key findings include that the number of phone calls per costumer as well as the number of unique calling partners per customer follow power law distributions. The authors also studied the evolution of their network, modeling its generative process with a double pareto LogNormal distribution. Last, Benevenuto et al. [2009b] studied silent interactions (browsing friends' pages) in online social networks. They found that the number of friends a user interacts with increases by an order of magnitude, compared to only considering visible interactions (explicit textual interaction).

Unlike all these studies, our analysis of YouTube focuses on characterizing *video* interactions in online social networks. One effort towards exploring multimedia interactions, particularly in the context of online education, is Wimba, which provides an audio message board service [Ross 2003]. However, a characterization of the properties of such interactions is not available. Thus, to the best of our knowledge, we provide the first effort towards understanding the properties that emerge from video-based user interactions.

In a previous work [Benevenuto et al. 2008a], we analyzed some properties of the social network created by video response interactions in YouTube. The present work greatly builds on this preliminary effort not only by providing a much more thorough, richer, and solid investigation of social network aspects of video interactions on YouTube, but also by characterizing several other aspects of the video response feature such as their popularity, duration, geographical origin, and other features.

3. CRAWLING A SOCIAL NETWORK

In order to analyze social network aspects on video interactions, we need to obtain information about users and their interactions (via video responses) from YouTube. To do it, we can sequentially visit pages on the YouTube site (that is, crawl) and gather information about YouTube video responses and their contributors. Every YouTube video post has a single contributor, who is a registered YouTube user. We say a YouTube video is a *responded video* or *video topic* if it has at least one video response. A video topic has a sequence of video responses listed chronologically in terms of when they are uploaded to the system.[7] We say a YouTube user is a *responded user* if at least one of her contributed videos is a responded video. Finally, we say that a YouTube user is a *responsive user* if she has posted at least one video response.

[6]www.cyworld.com.

[7]We note that YouTube does not allow a video to be posted as a response to more than one video topic.

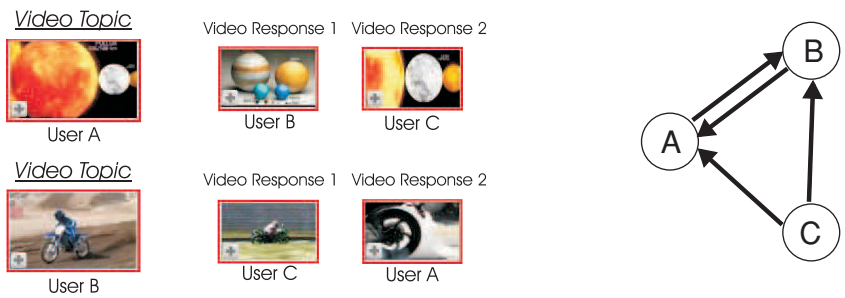

Fig. 1. Video responses posted to two video topics (left) and the graph created by these interactions (right).

A natural user graph emerges from video response interactions. At a given instant of time t, let X be the union of all responded users and responsive users. The set X is, of course, a subset of all YouTube users. We denote the *video response user graph* as the directed graph (X, Y), where (x_1, x_2) is a directed arc in Y if user $x_1 \in X$ has responded to at least one video contributed by user $x_2 \in X$. Figure 1 illustrates two video response sequences and the graph established by these interactions. We note that the video response user graph may have multiple weakly connected components.

Ideally, we would like to obtain the complete video response user graph (X, Y), which is equivalent to finding all the video responses and responded videos. Unfortunately, it is difficult to locate all video responses in YouTube. Although the YouTube site provides lists of the 100 most responded videos of all time, it does not currently provide a means to systematically visit all the responded videos. YouTube has numerous hyperlinking mechanisms from users to videos (e.g., users' favorite videos) and from videos to videos (e.g., related videos). But even by following all the hyperlinks among the videos and users, we cannot find all the users in X, since some of the users in X are not hyperlinked from any of the crawled pages. In particular, it is difficult to find the users in the small components of the graph (X, Y).

Instead, we design a sampling procedure that allows us to obtain a large representative subset A of X. For a given sampled subset A of X, let (A, B) be the directed graph, where (a_1, a_2) is a directed arc in B if user $a_1 \in A$ has responded to a video contributed by user $a_2 \in A$. Note that sampled graph (A, B) is a subgraph of (X, Y). It is desirable that the sampled set of users A have the following three properties:

—*Property 1*. Each connected component in (A, B) is a connected component in (X, Y), that is, the sampled subgraph (A, B) consists of entire connected components from (X, Y). As discussed in Mislove et al. [2007], this property is important in order to analyze the social networking interactions engendered by video responses.

—*Property 2*. The subset A covers a large fraction of X (at least 60%). Our sample would then include the majority of the responded and responsive users. Note that our interest is in crawling a sample of only those YouTube users who make use of the video response feature, which is a possibly much smaller subset of all YouTube users.

—*Property 3*. The most responded users are included in A. This last property ensures that we are including the most important users, and only neglecting users who have few responded videos.

In order to sample YouTube data following these three properties, we designed a distributed crawling framework similar to the one presented in Chau et al. [2007]. Our distributed crawler is composed of a master node and a number of slave nodes. The master node maintains a centralized list of users to

Algorithm 1. Video Response Crawler (run by slave nodes)

```
Input: A list L of users received from master node
for each User U in L do
    Collect U's information and video list;
    for each Video V in the video list do
        Collect information of V;
        if V is a responded video then
            Collect information of V's video responses;
            Insert the responsive users in list of new users NL;
        end if
        if V is a video response then
            Insert the responded user in list of new users NL;
        end if
    end for
end for
Return NL to the master node;
```

be crawled, which is initialized with a set of seeds. The master is also responsible for coordinating the operation of the slaves, sending nonoverlapping subsets of this list to them, thus preventing redundant crawling. The slaves, after obtaining the user identifiers from the master, crawl YouTube following Algorithm 1, return all new users collected to the master, and wait for a signal from it. The master, in turn, eliminates duplicate or previously crawled users from the received lists, and, in case there are still uncrawled users, starts a new round of crawling by sending new user identifiers to the slaves. Otherwise, the master sends a signal for termination to the slaves and stops execution. We note that the crawling process is terminated only after an entire weakest connected component of graph (X, Y) has been collected. We ran our crawler using 10 Linux boxes (1 master and 9 slaves) located at the Federal University of Minas Gerais in Brazil.

We used the set of all users who contributed the videos listed by YouTube as the all-time top-100 most responded videos[8] as seeds to our crawler. These consisted of 92 users. Our crawler produced the graph (A, B), composed of a large weakly connected component of graph (X, Y), thus, satisfying Property 1. The crawler ran for 5 days (September 21st–26th, 2007), gathering a total of 160,765 users, 223,851 responded videos and 417,759 video responses. For each video that was crawled, we collected a number of pieces of information available, including video identifier, video contributor identifier, title, category, description, tags, upload time, video duration, number of ratings, average rating, number of views, number of users who set the video as favorite, number of comments received, and number of video responses received. We also collected the author of the video responses of each video, and the sequence order in which the video responses were posted.

In order to verify Properties 2 and 3 of our graph (A, B), we ran Algorithm 1 using a random set of users as seeds. Algorithm 2 describes the mechanism used to randomly select users on YouTube. Out of the 4633 users collected with Algorithm 2, 3231 of them belong to A. Moreover, 67 out of the 100 randomly selected seed users are also in A. Both results are evidence that our sampling scheme may satisfy Property 2. In fact, our second crawl gathered a total of 182,725 users, out of which 146,799 can be found in our first data set. To verify Property 3, we ranked the 10, 100, and 1000 most responded

[8]YouTube usually provides lists of top videos or users with size 100.

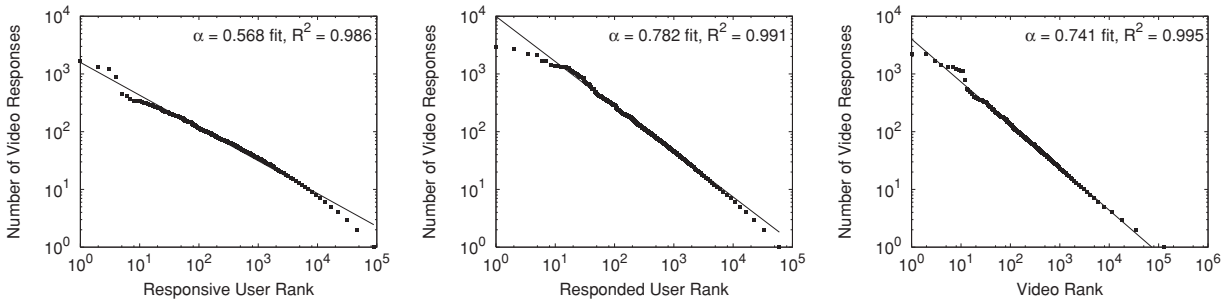

Fig. 2. Number of video responses per responsive user (left), responded user (middle) and responded video (right).

Algorithm 2. Selecting Random Seeds

Input: A list of words from a dictionary

```
Randomly select a word from the dictionary;
Search for a tag using YouTube API, using the selected word as tag;
for each Contributor C of the videos found in step 2 do
    if C is a responded OR a responsive user then
        Add user to the result list of users;
    end if
end for

Randomly select 100 users from the final list of users;
```

users from our second data set. We found that graph (A, B) contains all 10 of the 10 most responded users, 98 of the 100 most responded users, and 951 of 1000 most responded users. Thus, Property 3 is verified as well.

4. VIDEO RESPONSE STATISTICS

This section presents a characterization of several aspects of the video responses and responded videos collected, uncovering aspects of the usage of this feature by YouTube users.

4.1 Posting Characteristics

Experimental data have been used in the literature to show that the behavior of many real-world systems can be modeled by a power law distribution [Barford and Crovella 1998; Almeida et al. 1996; Breslau et al. 1999]. In terms of video interactions, we are interested in verifying the existence of the power law phenomenon to provide answers for three questions: (1) Do a small fraction of responsive users account for posting the majority of the video responses? (2) Do a small fraction of responded users receive the majority of the video responses? (3) Do a small fraction of responded videos receive the majority of the video responses?

Figure 2 (left) shows the distribution of the number of video responses posted by different users. We note that the 20% most responsive users contributed with 65% of all video responses, whereas 84% of all responsive users posted, each, less than five video responses. In fact, the distribution is well fitted by a power-law distribution (Prob(user with rank i posts a video response) $\propto 1/i^{\alpha}$) with $\alpha = 0.568$. Similarly, Figures 2 (middle) and (right) show the distributions of the numbers of video responses per responded user and responded video, respectively. Both distributions are also well fitted by power-laws with

Table I. Video Responses Across Categories

Category	Com	N&P	Ent	Mus	G&G	A&V	P&A	T&P	P&B	F&A	H&D	Spo
Comedy	**54.6**	2.8	10.1	6.6	1.8	0.5	1.4	0.8	13.8	5.5	1.1	0.9
News & Politics	5.3	**66.0**	4.4	4.0	0.4	0.4	0.6	1.2	13.5	2.1	1.7	0.5
Entertainment	8.8	2.1	**49.2**	10.1	3.9	0.5	0.8	1.1	10.4	10.3	1.1	1.5
Music	4.0	1.0	7.8	**72.6**	1.2	0.3	0.4	0.6	6.3	4.9	0.6	0.4
Gadgets & Games	2.8	0.3	6.0	3.3	**76.4**	0.4	0.1	0.2	4.3	5.2	0.7	0.4
Autos & Vehicles	4.1	2.9	6.8	1.9	4.3	**58.8**	0.4	4.3	6.1	5.7	2.2	2.5
Pets & Animals	6.4	1.9	4.3	2.9	0.4	0.2	**73.0**	1.4	6.1	1.9	1.0	0.6
Travel & Places	6.3	4.1	14.4	6.8	3.2	2.3	1.6	**38.9**	12.5	5.8	1.8	2.4
People & Blogs	5.6	4.6	6.8	5.3	1.1	0.4	0.9	1.3	**67.6**	3.3	2.4	0.8
Film & Animation	6.6	1.1	12.0	7.6	3.6	0.5	0.5	0.7	5.0	**61.2**	0.8	0.4
Howto & DIY	6.3	4.8	8.7	4.9	2.9	1.3	1.3	1.8	14.2	6.0	**46.6**	1.3
Sports	4.6	1.1	8.8	2.7	1.1	1.1	0.8	1.2	6.4	1.9	0.8	**69.4**

$\alpha = 0.782$ e $\alpha = 0.741$, respectively. To assess the accuracy of our proposed models, we measure the R^2 factor of the linear regression [Trivedi 2002] for each single distribution found. In our models, the values of R^2 are above 0.91 in all cases, which shows a good agreement ($R^2 = 1$ means perfect agreement).

In comparison with textual interactions, prior research showed that ranking blogs in terms of the number of comments received is also fitted by a power-law distribution with $\alpha = 0.7$ [Duarte et al. 2008].

4.2 Video Response Categorization

It is natural to expect that video responses posted to a video topic are related and approach the same subject. Next, we study the video categories of our collected data. The definition of a category is done by the users, contributors of the posted video, and the categories are chosen among 12 predefined subjects provided by YouTube.

Table I shows the distribution of video responses across categories. Each line refers to *responded* videos falling into one category, and reports the average fractions of their video responses falling in each category. The concentration of responses into different topics varies depending on the category of the responded video. For instance, most of the responses posted to "Gadgets & Games" videos are of the same category, whereas "Travel & Places" videos receive responses from different categories, with emphasis on "People & Blogs" and "Entertainment". Nevertheless, in most cases, the majority of the responses fall into the same category as the responded video. In other words, the responded video and most of its responses create a large group around the same subject. However, this subject is not strong enough to categorize all interactions triggered by a video, as a non-negligible fraction of the responses (24%–61%) falls outside the responded video category, creating multiple variable-size groups around each of the other categories.

Since users can freely assign predefined categories to their uploaded videos, we cannot assume that categories correctly describe the subject of the video content. However, some of the video responses might really be unrelated to the responded video. This, in turn may be a result of some sort of opportunistic behavior, such as spamming (see Section 7). Thus, information services such as advertising and recommendations should not take for granted that a video topic is directly related to its video responses.

4.3 Video Response Duration

Except for directors and other kinds of special accounts, YouTube users can post videos with a maximal duration of 10 minutes. Figure 3 (left) shows the distributions of video durations, considering separately, only responded videos and only video responses. Both distributions are very skewed, with 80% of all

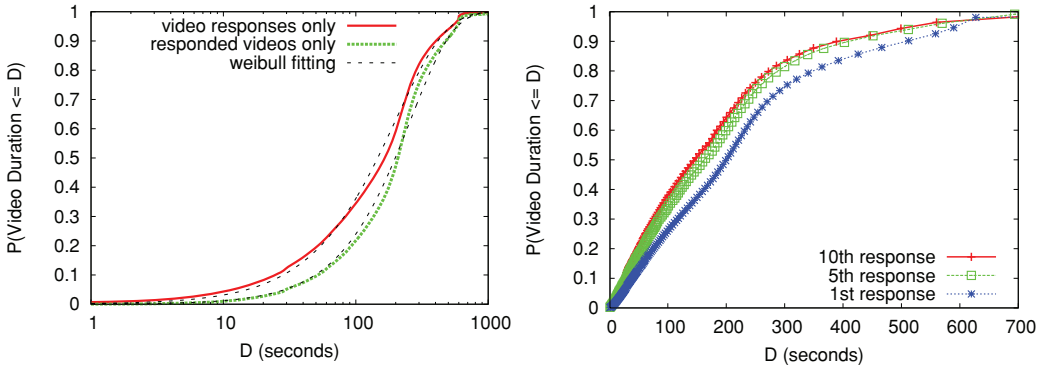

Fig. 3. Duration of responded videos and of (all) video responses (left), and of video responses at the same chronological position in the discussion thread (right).

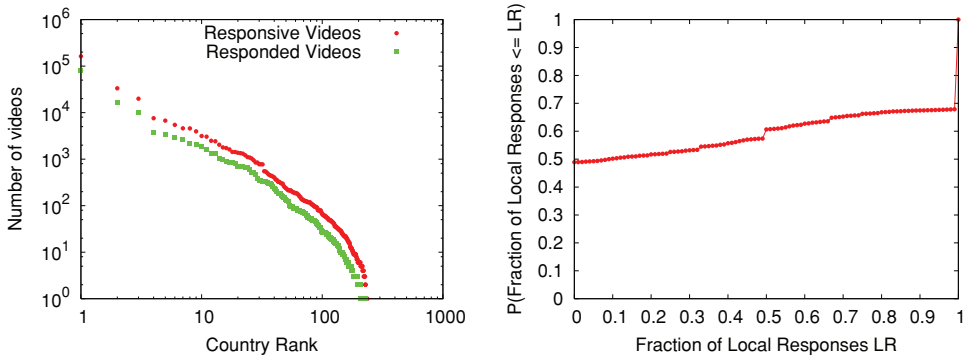

Fig. 4. Number of video responses per country (left) and percentage of local video responses (right).

samples being under 5 minutes. In fact, both follow Weibull distributions.[9] with parameters $\alpha = 0.0023$ and $\beta = 1.15$, for video responses, and $\alpha = 0.00054$ and $\beta = 1.35$ for responded videos. However, video responses have durations slightly more skewed towards shorter values.

Interestingly, we found that although the duration of a responded video has practically no impact on the number of responses it receives (correlation coefficient $C = -0.008$), there is a strong correlation between the duration of the responded video and the average duration of its responses ($C = 0.51$). Longer video topics tend to receive longer responses, and shorter video topics tend to receive shorter responses. Moreover, Figure 3 (right) shows that, considering only the ith responses of videos that had at least i responses, the duration distribution becomes more skewed as i increases. These results mimic the expected pattern in real-life human interactions, whereas longer (but interesting) expositions tend to initially trigger longer replies. However, replies tend to become shorter as the interaction progresses, and the discussion dies down.

4.4 Video Response Geographical Distribution

An important question that is often asked regarding Web characterization studies has to do with the geographical representativeness of the sample. Geography and friendship have been used to build real social network models [Liben-Nowell et al. 2005]. Figure 4 (left) shows the distribution of countries in

[9]The probability density function of the Weibull distribution is given by: $f(x) = (\beta/\alpha)^{-\beta}(x/\beta)^{(\beta-1)}e^{(-x/\alpha)^{\beta}}$.

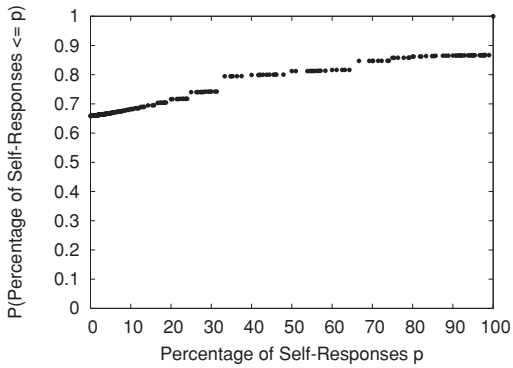

Fig. 5. Fraction of self-responses per responded video.

terms of number of videos shared in YouTube. The sample characterized is fairly diverse in terms of the number of countries, identified by the country described in the user profile, and of the numbers of video responses and of responded videos uploaded by users from different countries. Using the country identification, we are able to map the contributor population to over 236 countries. The top five countries in our sample account for almost 77% of all video responses uploaded to YouTube. The plots suggest a power law-like profile, with parameter $\alpha = 2.12\,(R^2 = 0.92)$ and $\alpha = 2.22\,(R^2 = 0.93)$ for video responses and responded videos, respectively. We also defined the percentage of local (from the same country) video responses as the ratio of video responses from the same country as the owner of the responded video to the total number of video responses. Figure 4 (right) shows the cumulative distribution of the percentage of local video responses over all responded videos of our sample. We notice that slightly more than half of the video conversations involve participants from the same country as the original contributor. For example, 40% of responded video has a percentage of local responses larger than 60%. This observation may be useful for designing geographically distributed video placement mechanisms.

4.5 Self-Responses

Interestingly, 25% of all video responses are self-responses, that is, responses posted by the user who posted the original video. Figure 5 shows the cumulative distribution of the fraction of self-responses posted to each responded video. Roughly 35% of the responded videos received at least one self-response, and around 12% of them received only self-responses.

Such a phenomenon was also observed in textual communication in Cyworld's guestbook, where 39% of all posts are self-posts [Chun et al. 2008]. In a guestbook, a self-post may have two purposes, namely a reply to a previous message and a note—a message to other users who access the guestbook. In the YouTube context, we conjecture that there are at least two purposes for self-responses in video interactions. Whereas some of these responses might actually be replies to other responses, others might be an attempt at self-promotion, that is, an attempt to inflate the number of video responses posted to a video to place it in a top position of the most-responded lists provided by YouTube, thus, gaining visibility in the system.

Therefore, a question that arises is whether a user can exploit the video response feature to raise the popularity of a video topic: is a video that receives many responses also viewed many times? The correlation between the number of responses and the number of views of a responded video shows this is not often the case (correlation coefficient $C = 0.16$). If we disregard all responded videos with at least one self-response, the correlation increases somewhat ($C = 0.24$), but remains low. These low correlation values indicate that one is not necessarily successful in artificially increasing the popularity

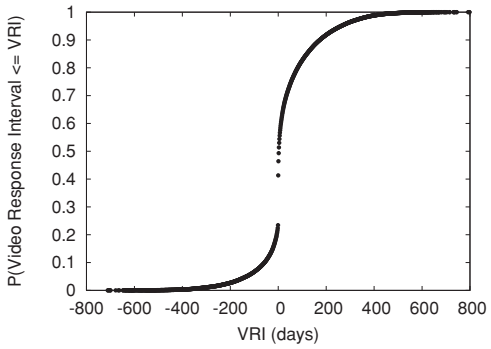

Fig. 6. Video response intervals (VRIs).

of a video by simply posting video responses to it. In other words, posting (self-)responses aiming at (self-)promotion does not necessarily pay off in YouTube.

4.6 Video Response Interval

A user may post a video response in one of three ways: (1) directly from the user's Web-cam; (2) choosing a video from one of the user's own, pre-existing YouTube contributions; (3) uploading a video from the user's disk drive. Unfortunately, YouTube does not provide a means to automatically determine in which manner a video was created. We thus propose to categorize video responses based on the time it was uploaded to YouTube, relative to the upload time of the responded video. We define the video-response-interval (VRI) as the upload time of the response minus the upload time of the responded video.

The cumulative distribution of the VRIs is shown in Figure 6. We can note that about 27% of the video responses correspond to videos uploaded before the responded video, and thus, were certainly not created as responses to it. This might be explained by the video content itself. For instance, we observed that one of the responded videos having many previously uploaded responses explicitly requested existing YouTube videos for responses. On the other hand, it might also be the result of users uploading existing (and not necessarily related) videos as self-responses. On the other hand, there are videos that clearly motivate users to post their opinions, creating long discussion threads. We found that 42% of the video responses are added within the first month after the video topic was uploaded, indicating a prompt reaction from responsive users. Nevertheless, a non-negligible fraction (17%) of responses are added long after the video appeared in the system (VRI $\geq$ 100 days), meaning that some videos exhibit long-term popularity, and new interactions are initiated long after they were uploaded.

4.7 Summary

In this section we made several key observations, summarized as follows.

—The distributions of video responses across responsive and responded users and responded videos follow power laws. This means that a small number of responded and responsive users account for receiving and posting a large number of video responses. Moreover, a large number of video responses are posted to a small set of very popular video topics.

—A responded video and most of its responses are typically associated with the same category (e.g., "News&Politics", "Music", and so on). However, a significant fraction of the video responses (24%–61%) fall outside this group, covering each other category available in YouTube.

—The durations of responded videos and of video responses both follow Weibull distributions, although the duration of video responses are more skewed towards shorter values. Moreover, there is a strong

correlation between the duration of the responded videos and the average duration of their respective responses.

—Around 40% of the responded videos have a fraction of responses originated from the same country superior to 60%, an observation that might be exploited in the design of geographically distributed video placement mechanisms.

—Around 25% of all video responses are self-responses, a phenomenon also observed in textual communications in the Cyworld's guestbook, but with a higher frequency, as 39% of all posts are self-posts. Moreover, 35% of the responded videos receive at least one self-response, that is a video response posted by the user who posted the responded video. However, the correlation between the number of video responses and the number of views of a responded video is weak.

—A significant fraction (27%) of the video responses posted to a video topic are actually uploaded to YouTube before the responded video. Nevertheless, about 42% of all video responses are posted within one month after the video topic, thus indicating prompt responsive users. In contrast, 17% of the video resposts are posted at least 100 days after the topic, a sign of long-term popularity.

5. VIDEO-BASED INTERACTIONS

In order to understand the characteristics of video-based interactions, we adopt a hierarchical approach to analyze the collected data, looking at the results of the crawls as a two-level hierarchy. This section focuses on the lower level, that is, on user interactions engendered by a single responded video. The next section explores the top-level interactions, namely, the social network created by the interactions among users across all responded videos.

Our characterization of user interactions at the video level consists of examining the sequence of video responses following a YouTube video topic. For each responded video V_i, let n_i denote the number of video responses of V_i, and let us denote the sequence of V_i's video responses by $\{VR_{i,1}; VR_{i,2}; \ldots ; VR_{i,j}; \ldots ; VR_{i,n_i}\}$, where $VR_{i,j}$ is the jth video response of video V_i.

A user may add multiple video responses in sequence to the same video. We define a sequence of responses $S_{i,k}$ as a series of *consecutive* responses from the same user to video V_i. Thus, the ordered list of video responses to video i can be also expressed as $\{S_{i,1}; S_{i,2}; \ldots S_{i,s_i}\}$, $1 \leq s_i \leq n_i$, where $S_{i,k}$ is the sequence $\{VR_{i,j}; VR_{i,j+1} \ldots\}$ of *consecutive* responses from user U_k to video V_i. Note that, the same user may post multiple (non-consecutive) sequences of responses to the same video V_i.

In the following example, video V_i received 7 video responses grouped into 4 sequences, posted by 3 different users (U_1, U_2 and U_3):

$$V_i; \overbrace{\underbrace{VR_{i,1}, VR_{i,2}}_{U_1}}^{S_{i,1}}, \overbrace{\underbrace{VR_{i,3}}_{U_2}}^{S_{i,2}}, \overbrace{\underbrace{VR_{i,4}, VR_{i,5}, VR_{i,6}}_{U_1}}^{S_{i,3}}, \overbrace{\underbrace{VR_{i,7}}_{U_3}}^{S_{i,4}}.$$

Our characterization focuses on a simple metric, defined as the ratio U-S of the *Number of Unique Responsive Users* to the *Number of Sequences of Responses*. In this example, the ratio is 3/4. A video-based interaction with a ratio U-S close to 0 indicates an asynchronous video dialogue between a relatively small number of highly active users, who keep the discussion alive with multiple (not necessarily consecutive) responses to each other. This type of interaction is akin to the exchanges and debates in a parlor or public forum, in which the communication underscores a many-to-many dialogue among participants. At the other extreme, when the ratio approaches 1, there may be two types of interaction. One type occurs when the number of unique responsive users equals the number of sequences of responses, resembling a register, petition or guestbook, for which the communication is many-to-one, and the purpose of a video response is to record a comment (or support a petition, etc.). The other type has just one

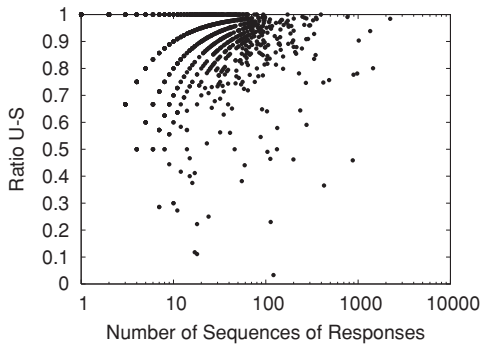

Fig. 7. Types of video-based interaction.

user posting a single sequence. Single-user sequences containing multiple responses are typically an attempt at promotion.

Figure 7 shows the ratio U-S versus the number of sequences of responses for each responded video in our data set. Different types of interactions can be seen across all responded videos, characterized by a wide range of U-S values. However, only a relatively small number of videos triggered lively interactions, with each responsive user adding on average at least 3 sequences of responses. In fact, the vast majority (99%) of the responded videos triggered interactions with only one sequence of responses per responsive user (ratio equal to 1). In fact, Figure 7 shows that these interactions occur among groups of responsive users of varying sizes (number of sequences). We further analyze this issue by looking into the total number of responses included in each sequence. We found that only 6% of all sequences have more than one video response. We also found a few cases in which the responded video received a single sequence with multiple responses: 714 videos received a single sequence with at least 5 responses from the same user, whereas 146 videos had single sequences with 10 or more responses from the same user. Finally, we found that 13% of all video topics with more than two video responses received only self-responses.

In summary, in this section we studied the patterns of user interactions engendered by a single responded video. We showed that the vast majority (99%) of the responded videos triggered interactions for which each user participated only once, resembling what occurs in a guestbook. Nevertheless, we also found that a few videos triggered very lively interactions, with each responsive user participating at least three times.

6. NETWORK-LEVEL INTERACTIONS

A social network is a set of people or groups of people with some pattern of interactions between them [Newman and Park 2003]. Social networks are useful for analyzing social phenomena that involve the interactions of a large number of heterogeneous entities, such as videos, contributors, and viewers. In this section, we study the video response user graph (A, B), defined in Section 3. Table II presents the main statistics of the network built from the graph (A, B) and its largest strongly connected component (SCC).

6.1 Node Degree

The key characteristics of the structure of a directed network are the in-degree (k_{in}) and the out-degree (k_{out}) distributions. As shown in Figure 8, the distributions of the degrees for the entire graph follow power laws $P(k_{in/out}) \propto 1/k_{in/out}^{\alpha_{in/out}}$. The scaling exponents of the whole network lie in a range of 2.0 and

Table II.
Summary of the Network Metrics (CV = Coefficient of Variation)

Characteristic	Entire Network	Largest SCC
Number of Nodes	160,074[11]	7,776
Number of Edges	244,040	33,682
Average Clustering Coefficient	0.047	0.137
Number of Nodes of Largest SCC	7,776	7,776
Number of Components	149,779	1
Assortativity/Pearson Coefficient r	−0.017	0.017
Average Distance	8.40	8.40
Average in-degree k_{in} (CV)	1.53 (9.38)	4.33 (3.14)
Average out-degree k_{out} (CV)	1.53 (1.717)	4.33 (1.28)

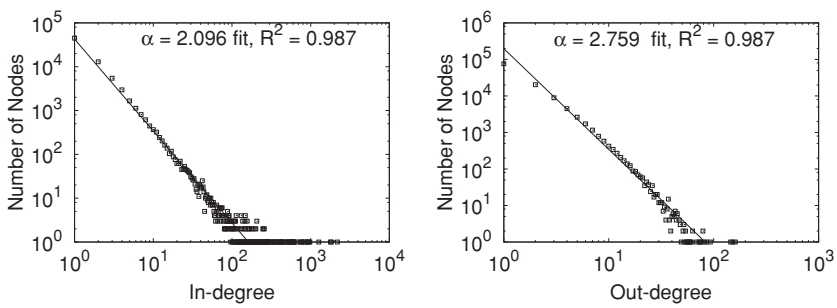

Fig. 8. In-degree and out-degree distributions (entire network).

3.4, which is a very common range for social and communication networks [Ebel et al. 2002]. Our results agree with previous measurements of many real-world networks that exhibit power law distributions, including the Web and social networks defined by blog-to-blog links [Kumar et al. 2003] as well as the discussion threads in Slashdot [Gómez et al. 2008].

The in-degree exponent is smaller than the exponent of the out-degree distribution, indicating that there are more users with larger in-degree than out-degree. This fact suggests a link assymmetry in the directed interaction network. Unlike other social networks that exhibit a significant degree of symmetry [Mislove et al. 2007], the user interaction network shows a structure similar to the Web graph, where pages with high in-degree tend to be authorities and pages with high out-degree act as hubs directing users to recommended pages [Kleinberg 1999]. In order to investigate this point further, Figure 9 shows the cumulative distribution of the ratios of in-degree to out-degree and out-degree to in-degree for the user interaction network. The network has 60% of the users with out-degree higher than in-degree and 5% of the users with significantly higher in-degree than out-degree (ratio > 3). This is evidence that a few users act as "authorities" and "hubs". We have observed in our dataset that authority-like users (that is, highly responded users), with high in-degree, are typically media companies that upload professional content, including sports, entertainment video, and TV series. Nodes with very high out-degree may indicate either very active users or spammers: users that distribute unsolicited content to legitimate users.

According to Newman and Park [2003], assortative mixing is a graph theoretical quantity, typical of social networks. We then investigate this structural property in the user interaction network. A

[11]The number of nodes on the graph is smaller than the number of users collected because we excluded users with only self-responses from the graph.

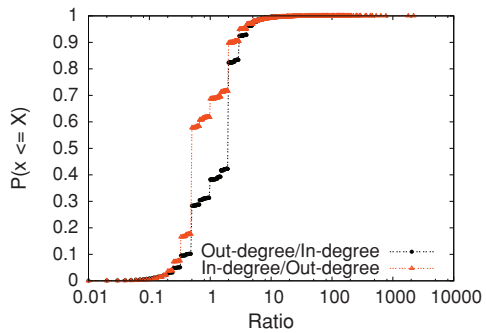

Fig. 9. In-degree and out-degree ratios.

network is said to exhibit assortative mixing if the nodes with many connections tend to be connected to other nodes with many connections. Social networks usually show assortative mixing. The assortative (or dissortative) mixing of a network is evaluated by the Pearson coefficient r, which is calculated as follows [Newman 2002]:

$$r = \frac{\sum_i j_i k_i - M^{-1} \sum_i j_i \sum_{i'} k_{i'}}{\sqrt{\left[\sum_i j_i^2 - M^{-1}\left(\sum_i j_i\right)^2\right]\left[\sum_i k_i^2 - M^{-1}\left(\sum_i k_i\right)^2\right]}}, \tag{1}$$

where j_i and k_i are the excess in-degree and out-degree of the vertices that the ith edge leads into and out of, respectively, and M is the total number of edges in the graph. Positive values of r indicate that the network exhibits assortativity—nodes with high degrees tend to be connected to nodes with high degrees. On the other hand, negative values of r indicate that the network exhibits a dissortative mixing—nodes with high degrees tend to be connected to nodes with low degrees.

Table II shows values of r for the directed graph of the interaction network. The video response user graph exhibits a dissortative mixing with $r = -0.017$, where high degree nodes preferentially connect with low degree ones and vice versa. Analyzing only the largest SCC, we can observe an assortative mixing $r = 0.017$. The existence of a significant assortative mixing is associated with the notion of social communities [Newman and Park 2003]. Thus, whereas the entire user interaction graph does not show evidence of formation of a social community, its largest SCC exhibits some sign (though weak) of such behavior.

6.2 Clustering Analysis

It has been suggested in the literature that social networks possess a topological structure where nodes are organized into communities [Newman and Park 2003], a feature that can account for large values of clustering coefficient and degree correlation. The clustering coefficient of a node i, $cc(i)$ is the ratio of the number of existing edges between i's neighbors over the number of all possible edges between them.

The clustering coefficient of a network, CC, is the mean clustering coefficient of all nodes. The average CC of all nodes of our network is $CC = 0.047$, whereas the mean clustering coefficient for a random graph with identical degree distribution and number of nodes but random links is $CC = 0.007$. Thus, our network presents a significantly larger clustering coefficient than a randomized version of the graph, which points to the existence of some structural hierarchy and the presence of small communities in the graph.

Figure 10 (left) shows the cumulative distribution of the node clustering coefficient. The network contains a significant fraction of nodes with clustering coefficient equal to zero. Specifically, 80% of all

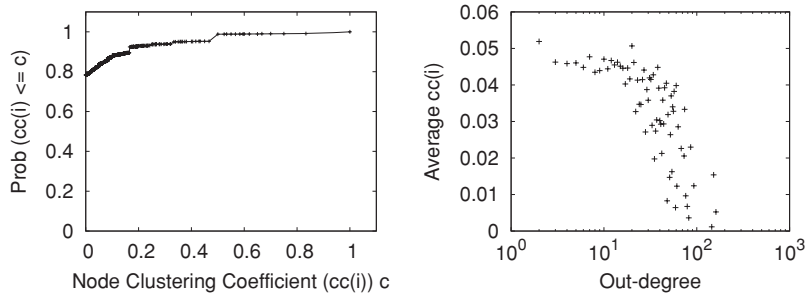

Fig. 10. Distribution of node clustering coefficient $cc(i)$ (left) and average $cc(i)$ per out-degree (right).

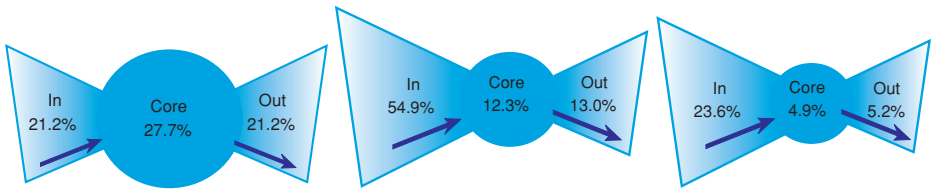

Fig. 11. Bow tie structures of the Web (left), Java Forum (middle), and the video response user graph (right).

nodes in the entire user interaction network have $CC = 0$. This feature indicates that there is a clear difference between average clustering in the entire network and in individual nodes. Figure 10 (right) shows how the clustering coefficient varies with the node out-degree. Higher values of the clustering coefficient occur among low out-degree nodes, suggesting the lack of large communities around high out-degree nodes.

6.3 Component Analysis

Since not all the users in the network are both responsive and responded, we adopted a model called *bow tie structure* to study our network structure. The bow tie structure was initially proposed to provide insights of the organization of the network structure of the Web [Broder et al. 2000a]. The key idea is that the Web graph (nodes are Web pages and edges are links between the pages) has distinct components depicted in a figure that resembles a bow tie. In our context, the central core contains users who frequently communicate via video interactions by posting and receiving video responses. It is the largest strongly connected component (SCC), meaning that one user can reach every other user following video response links. The *in* component contains users who post video responses to the core, but cannot be reached from it. The *out* component consists of users who receive video responses posted by users in the core but do not link to it, such as directors or companies specialized in providing video content for YouTube. Other users (the Tendrils and Tubes), connect to either the in or out components, or both, but not to the core. They are users who receive video responses from the in users or who post video responses to the out users. Tendrils and Tubes are not represented in our analysis.

Figure 11 compares the bow tie structures of the Web [Broder et al. 2000a], of a textual Java Forum [Zhang et al. 2007] (users are nodes and a directed edge from A to B means that A had a question answered by B), and of our video response user graph. Analyzing the fractions of nodes in the three major components, shown in the figures, we make two interesting observations. First, the bow tie structure of the video response interactions is very different from that of the Web, but shares some characteristics of the Java Forum structure. In particular, the fraction of nodes in each component is much more balanced in the Web structure, whereas there is a clear unbalance towards the in component

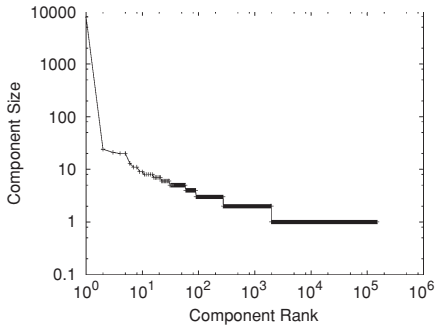

Fig. 12. Component size rank.

in the other two. Moreover, if we divide the fraction of users falling into each component of the Java Forum structure by the corresponding fraction in the video response structure, we get ratios equal to 2.3, 2.5, and 2.5, for the in, out, and core component, respectively, indicating a high structural similarity of the video response interactions with those in the textual forum.

Second, in both the video response user graph and in the Java Forum graph, there is a large number of users who post video responses and interact with the core users (users from the in and core components). However, only a small fraction of these users receive responses back. In fact, the video response user graph has the smallest core and out components, among the three structures.

There are two particular characteristics of the YouTube video response feature that might impact the video response graph bow tie structure. The first is the fact that a video can receive multiple video responses, but it can be posted as video response to only a single video. This restriction of the YouTube video response feature allows the in component to grow freely, but limits the growth of the core component by imposing restrictions to the creation of response back-links. The second is that the community structure developed by video interactions might be at an early stage. In fact, in order to further investigate this, we did another crawl following the same methodology presented in Section 3. The crawl was executed from 01/11/2008 to 01/18/2008. The core component of the graph created from this data set has a size of 5.02%, which is slightly larger than the core of our previous data set. On the other hand, the in component in the new data set is starting to diminish (22.08%). It seems that some new paths from the core to nodes in the in component have been established, thus increasing the size of the core component.

Next, we investigate the network properties of the largest SCC, which corresponds to the core component, of our video response user graph. In spite of its its small size, the core component concentrates 10% of the views and 22% of the video responses, and thus, deserves further analysis. Our main findings are summarized in Table II. In particular, the average clustering coefficient of the largest SCC is equal to 0.137, three times larger than the clustering coefficient of the entire network. Thus, user interactions captured by the largest SCC form a much more tightly connected community.

Figure 12 shows the distribution of the sizes of all strongly connected components of our network, sorted from the largest component to the smallest one. The core component is significantly larger than the others. The distribution suggests a general structure that includes the largest SCC, a large number (147,805) of components with a single node, and a number of variable (small) sized components (1974 components, in total). As we are working with a directed graph, the components with size one (singletons) are nodes with links in only one direction. The other smaller components are groups of users who represent small sized communities, composed of a few people like a family or a group of friends, who express their interests and establish communication via video responses. In fact, there is a

significantly negative correlation between the size and the average clustering coefficient of the components (correlation $C = -0.36$), indicating that small components tend to be more tightly connected.

6.4 Link Reciprocity

Another interesting metric to observe is the link reciprocity coefficient ρ, a metric that captures the reciprocity of the interactions in the entire network [Garlaschelli and Loffredo 2004]. Given the adjacency matrix of a directed graph, that is $a_{ij} = 1$ if there is a link from i to j, and $a_{ij} = 0$ otherwise, the reciprocity coefficient ρ is defined as the correlation coefficient between the matrix entries. In other words:

$$\rho = \frac{\sum_{i \neq j}(a_{ij} - \bar{a})(a_{ji} - \bar{a})}{\sum_{i \neq j}(a_{ij} - \bar{a})^2}, \tag{2}$$

where the average value $\bar{a} = \sum_{i \neq j} a_{ij}/N(N-1)$, and N is the number of users in the graph.

The reciprocity coefficient tells whether the number of mutual links in the network is more or less than that of a random network. If the value of ρ is higher than 0, the network is reciprocal; otherwise, antireciprocal. The ρ value for the video response user graph is 0.051, which means that the network is more reciprocal than a randomized graph. In contrast, the reciprocity coefficients of the World Wide Web and of Wikipedia[12] have been found to be 0.5161 [Garlaschelli and Loffredo 2004] and 0.32 [Zlatic et al. 2006], respectively. In terms of networks created essentially by textual communication, the reciprocity coefficient is 0.231 for e-mail [Garlaschelli and Loffredo 2004], 0.28 for Slashdot[13] [Gómez et al. 2008], 0.58 for Twitter [Java et al. 2007], and 0.765 for guestbook communication in Cyworld [Chun et al. 2008].

Thus, among all these networks the quantitative link reciprocity of the video response communication is the smallest. In conclusion, our analysis shows that the video response feature triggers weakly reciprocal communication, specially in comparison with networks built from textual interactions.

6.5 Average Distance

Last, we look at the average distance of our video response network, which is the average number of hops along the shortest paths for all possible pairs of network nodes [Newman and Park 2003]. The average distance of our network is 8.40, a low value if compared with other directed graphs such as the Web (average distance equal to 16.12) [Broder et al. 2000a]. As shown in Mislove et al. [2007], the average path length between nodes is usually short in social networks.

Such a small average distance combined with an average clustering coefficient almost 7 times larger than that of the same graph with randomized links (see Section 6.2) constitute the main properties of a small world graph [Albert et al. 1999; Broder et al. 2000b]. These properties were also verified in textual communication contexts such as MSN instant messages [Leskovec and Horvitz 2008] and in friendship relations in LiveJournal, Flickr, Orkut, and YouTube [Mislove et al. 2007].

6.6 Summary

We summarize the main findings of the network-level properties of the video response user graph as follows.

—In-degrees and out-degrees follow power law distributions, and the network exhibits small world properties.

[12]www.wikipedia.com.
[13]www.slashdot.com.

—The community structure contains a reasonably small strongly connected component (SCC) and a large number of much smaller components (up to 20 members) and of singletons. Small communities tend to be more tightly connected than large components, but users on the largest SCC are more tightly connected in comparison with the whole network.

—The community structure developed by video responses seems to be at an early stage, as the size of the SCC, though small, seems to be increasing.

—In comparison to previously characterized textual communication networks, the video response user graph shares similar degree distributions, small world properties, as well as bow tie structures. However, in sharp contrast, the degree of reciprocity in video-based interactions is significantly weaker.

7. OPPORTUNISTIC BEHAVIOR

Different forms of unsolicited communication are taking a toll on users of social networking services [Zinman and Donath 2007]. Unsolicited communication, here referred to as opportunistic behavior, opens a large gray area, where videos could be considered spam or promotion. One form of spam occurs when users submit a video with a long list of misleading tags to describe its content in order to fool video searching mechanisms [Heymann et al. 2007]. Another form of video spam occurs when a video is posted as a response to a video topic, but whose content is completely unrelated to the video topic [Benevenuto et al. 2008b]. On the other hand, promotion consists of users trying to boost the ranking of a video to make it more visible in the social network ranks. Towards that goal, users upload a large number of video responses to a video topic, most of which are not necessarily related to the responded video.

Due to its intrinsic nature, video response appears to be an attractive feature to users willing to spam in order to promote specific content, advertise to generate sales, disseminate pornography (often as an advertisement) or simply compromise the system reputation. Unlike textual responses or comments, one has to actually watch the video to realize it is, or contains, some form of spam or promotion, consuming system resources, in particular bandwidth, and compromising user patience and satisfaction with the system.

A simple example illustrates how opportunistic behavior, such as spamming, promotion, and unexpected advertising, could be identified by examining user behavior patterns in a social networking service. Figure 2 (left) shows that the distribution of the number of video responses posted by responsive users follows a power law. Experimental data have been used in the literature to show that the behavior of many real-world systems can be modeled by a power law distribution. We notice from the figure that three points (leftmost part of the curve) do not fit well the expected power law. They represent users who have posted a much larger number of video responses than predicted by the model. By identifying these users in our dataset, we realized they share common characteristics, namely: (1) they post video responses to either their own videos or to other specific videos to boost their rankings in order to increase their visibility, and (2) they make use of video responses as a marketing opportunity in the social networking environment, spreading video-based information to influence others or to advertise commercial products and services.

In the following, we discuss two simple metrics that may help understanding different types of user behaviors in video-based interactions, and may serve as a first cut in the identification of opportunistic behavior.

7.1 User Inter-Reference Distance

We define the Inter-Reference Distance (IRD) of a user who uploads video responses to a video topic as the sum of the numbers of responses (plus one) that appear between two consecutive video responses

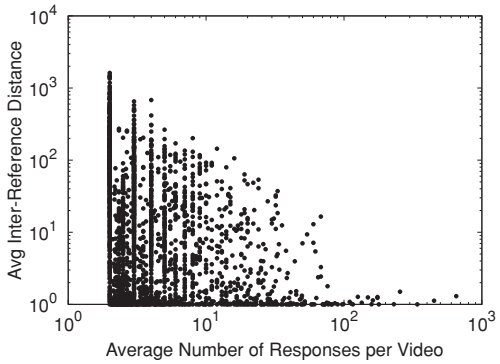

Fig. 13. Temporal patterns of user participation in sequences of video responses.

of the same user. As an example, consider the IRDs computed to the following two video topics and the contributors of their video responses.

Video Topic 1: $U_1\ U_2\ U_2\ U_1$ $\quad$ $\mathrm{IRD}(U_1) = 3$, $\mathrm{IRD}(U_2) = 1$
Video Topic 2: $U_1\ U_1\ U_1$ $\quad$ $\mathrm{IRD}(U_1) = 1 + 1 = 2$

We compute a user's average IRD by first calculating her IRD for each video she responded to, and then taking the average over all videos responded by her. In this example, the average IRD for user U_1 is computed as $(3 + 2)/2 = 2.5$.

The IRD metric allows us to assess temporal patterns of the users' participation in a sequence of video responses. Previous studies on spam characterization refer to the importance of analyzing temporal issues in order to detect malicious and opportunistic behavior [Gomes et al. 2007]. For example, whereas traditional e-mail traffic is concentrated on diurnal periods, the arrival rate of e-mail spams is roughly stable over time [Gomes et al. 2007].

Figure 13 plots the average IRD for each responsive user as a function of the average number of video responses per video responded by the user. A user who uploads many video responses per video, one after the other, mostly following a mechanical process, might be a candidate for further investigation. Thus, the combination of a large number of video responses per video and a small average IRD suggests the user exihibits some type of opportunistic behavior, such as spamming.

We conducted an investigation to verify if the combination of these two metrics (average IRD and average number of responses per responded video) could accurately be used to identify users with opportunistic behavior. Our experiment focused on users who are located on the rightmost part of Figure 13. In the dataset, a total of 298 users have average IRD less than 3 and average number of responses per video greater than 10. A group of volunteers in our laboratory randomly selected 95 from these 298 users to be analyzed.[14] Our volunteers viewed the video responses posted by each user, and, according to their content, classified the responsive user into either social or opportunistic. If at least one video response was considered spam or promotion, the responsive user was labeled as opportunistic. A total of 80% of the analyzed users were classified as opportunistic, suggesting that the proposed metrics could be a starting point to develop heuristics to combat opportunistic behavior in video interactions.

[14]The sample size was chosen so as to guarantee a 90% confidence level and errors under 7% [Jain 1991].

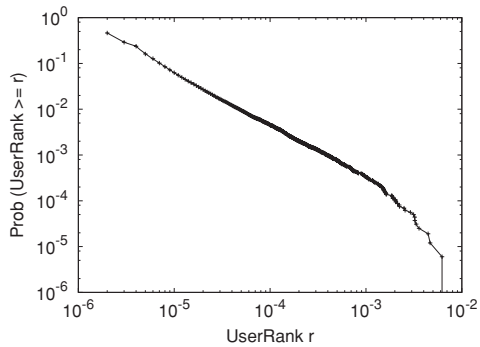

Fig. 14. UserRank scores.

7.2 User Rank

The next step is to use the structure of the social network for detecting opportunistic patterns. Towards that goal, we explore the use of the PageRank algorithm [Brin and Page 1998], first proposed for the Web context. The intuition behind PageRank is that a Web page is important if it has many incoming links or if the page has links coming from highly ranked pages. Thus, the importance of a certain Web page influences and is influenced by the importance of some other pages.

We explore the application of the PageRank algorithm to determine the importance of a user in the video communication network. In this context, we refer to the scores computed by the algorithm as UserRanks, which represent indicators of the importance of users in terms of their participation in video interactions. More formally, consider our video response user graph $G = (A, B)$ consisting of a set A of N users (vertices) and a set B of directed links (edges) that connect users interacting via video responses. The UserRank score $UserRank(U_p)$ of a user $U_p \in A$ is defined as:

$$UserRank(U_p) = d \sum_{U_q:(U_q,U_p)\in B} \frac{UserRank(U_q)}{\omega(U_q)} + (1-d)\frac{1}{N}, \tag{3}$$

where d is a damping factor and $\omega(U_q)$ is the out-degree of user U_q. Thus, the score of some user U_p is a sum of two components, one accounting for the scores of users pointing to U_p, and the other corresponding to a fixed value depending only on the size N of the network. In the following, we set $d = 0.85$, as commonly used [Brin and Page 1998].

Figure 14 shows the complementary cumulative distribution of the UserRank scores in our network. Similarly to the Web [Pandurangan et al. 2002], it clearly follows a power law, with most users with very low scores. However, there are a few users with very high scores. We analyzed the profiles and the video responses of a few users in both extremes of the curve. Users with high scores are among the most responded and viewed. Most of them are directors (a director account has special privileges in YouTube). In comparison, users with low scores own videos that are rarely viewed and that receive at most only a few video responses from the video community.

In order to assess if the UserRank scores capture the importance of a user in the video response user graph, we computed the correlation coefficient C between the UserRank and several other characteristics of users and of their videos. As shown in Table III, there exists a strong correlation between UserRank and two metrics, namely the total number of ratings received by the user's videos, and total number of views of the user's videos, suggesting that the algorithm captures the importance of users in terms of the numbers of views and of ratings of their contributions. These strong correlations can also be seen in the scatter plots presented in Figure 15. In contrast, the correlations between UserRank scores

Table III. Correlation Coefficient of the UserRank with Different Characteristics of Users and their Videos

Characteristic	Correlation Coefficient
Total Number of Ratings Received by User's Videos	0.44
Total Number of Views of User's Videos	0.27
Total Number of User's Videos Favorited	0.17
Link Reciprocity	0.14
Out-degree	0.13
Number of Friends	0.11
Number of Videos	0.07
Clustering Coefficient	0.04

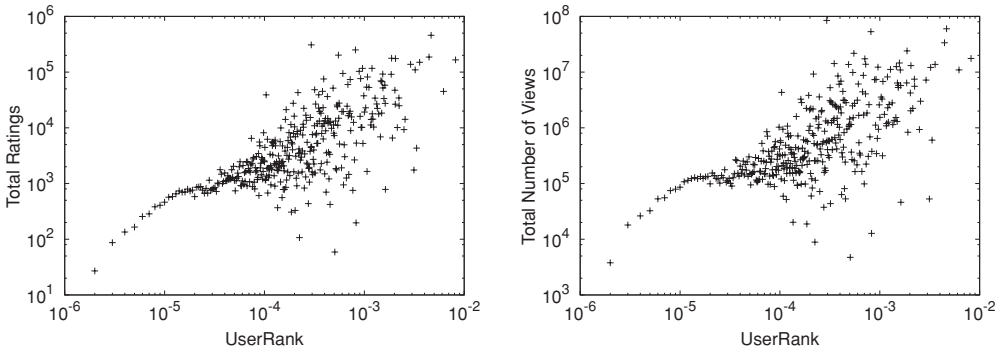

Fig. 15. Correlation between UserRank and total ratings (left), and total number of views (right).

and all other analyzed metrics such as link reciprocity, number of friends, number of videos uploaded, and clustering coefficient, are much weaker.

Next, we assess the potential benefit of using the UserRank score to detect users that exploit the video response feature to boost video ranks, aiming at promoting their content. Intuitively, boosting a video rank can bring some extra visibility it. However, it may also be perceived as a spam or undesired content by other users. Moreover, videos that quickly reach a high rank are strong candidates to be kept in caches or in content distribution networks (CDNs) [Cha et al. 2007]. Thus, promoted videos can be confused with popular videos, impacting not only the user satisfaction with the system, but also the performance and scalability of online video social network services.

Our strategy to detect suspect users consists of searching for users with low UserRank scores who have videos on the top lists. The intuition behind it is the following. Suppose a video had its rank boosted by a series of self-responses or video responses posted by possibly fake accounts with low rank. Suppose also, that the number of (self-)responses posted was such that the video made it to one of the top-100 lists of most responded videos kept by YouTube. Clearly, the UserRank score of the contributor of this promoted video should be lower than the scores of the contributors of the other non-promoted videos in the same top-100 list, as these received video responses from several different users, possibly with different scores.

In order to verify if the UserRank can be useful to pinpoint potential promoters among the contributors of the most responded videos, we conducted the following experiment. We first took the videos in the all-time top-100 most responded video list (used as seeds to our crawler), and sorted them in increasing order according to the video contributor's UserRank. Next, we manually inspected each video of this sorted list, starting from its beginning, and determined, after each round of 10 videos inspected, the total number of promoted videos found. Figure 16 illustrates this experiment, showing the cumulative

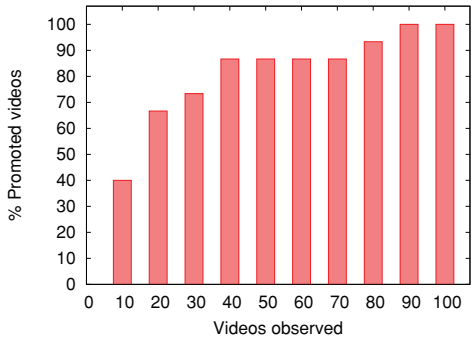

Fig. 16. Detecting promoted videos in the top-100 most responded list.

percentage of promoted videos found after each round of the experiment. By observing only 40 videos of the users with the lowest UserRank scores, we are able to identify 87% of the promoted videos in the top-100 most responded list, a much higher fraction than what we would have achieved if we had randomly selected the videos to observe.

7.3 Summary

User-generated video systems are amenable to different types of opportunistic behaviors such as promotion, unsolicited product advertising, and video spamming In this section we presented evidence of these opportunistic behaviors in the use of the video response feature in YouTube. We also proposed the use of two metrics, namely, the Inter-Reference Distance (IRD) and the UserRank, to help detect opportunistic behavior. An initial assessment indicates that:

—Combining the user IRD with the average number of responses posted by her to each video may help identify opportunistic users who exploit the video response feature to do spamming or promotion.

—The UserRank scores may be used to detect contributors of promoted videos among the top lists in YouTube.

8. CONCLUSIONS AND FUTURE WORK

Online video social systems, such as YouTube, provide a video response feature that allows users to post a video as a response to another. Such a feature enriches the online interaction among users, allowing them to exchange knowledge and express ideas through video interactions.

In this work, we performed an extensive characterization of users interacting with each other through the YouTube video response feature. In addition to providing statistical models for various characteristics, our study unveils interesting findings about video-based communication and shows evidence of opportunistic behavior. More importantly, the characteristics of video interactions that we presented are significantly different from those of traditional computer-based interactions with text and image. Part of the difference stems from the change from textual communication to video-based communication, creating a new paradigm for online communication.

We believe that the opportunistic behavior patterns we unveiled may jeopardize the trust of users on the system, compromising its success in promoting social interactions. Thus, our current work is focused on using social network aspects and video characteristics to identify opportunistic users in online video social systems. In fact, we recently approached the problem of identifying promoters and video spammers by applying machine learning techniques [Benevenuto et al. 2009a]. As future work, we plan to address the identification of other kinds of opportunistic behavior, such as bad association of

tags to videos in an attempt to fool video search engines (tag spamming). Another interesting direction is to explore the causes for a video to be interesting, triggering video response conversations, as recently studied for textual comments in YouTube [Choudhury et al. 2009].

REFERENCES

ALBERT, R., JEONG, H., AND BARABASI, A. 1999. The diameter of the World Wide Web. *Nature 401*, 130–131.

ALMEIDA, V., BESTAVROS, A., CROVELLA, M., AND OLIVEIRA, A. 1996. Characterizing reference locality in the WWW. In *Proceedings of the International Conference on Parallel and Distributed Information Systems (PDIS'96)*.

BARFORD, P. AND CROVELLA, M. 1998. Generating representative Web workloads for network and server performance evaluation. *SIGMETRICS Perform. Eval. Rev. 26*, 1, 151–160.

BENEVENUTO, F., DUARTE, F., RODRIGUES, T., ALMEIDA, V., ALMEIDA, J., AND ROSS, K. 2008a. Understanding video interactions in Youtube. In *Proceedings of the ACM International Conference on Multimedia (MM)*.

BENEVENUTO, F., RODRIGUES, T., ALMEIDA, V., ALMEIDA, J., AND GONALVES, M. 2009a. Detecting spammers and content promoters in online video social networks. In *Proceedings of the International ACM SIGIR Conference*.

BENEVENUTO, F., RODRIGUES, T., ALMEIDA, V., ALMEIDA, J., ZHANG, C., AND ROSS, K. 2008b. Identifying video spammers in online social networks. In *Proceedings of the International Workshop on Adversarial Information Retrieval on the Web (AIRWeb)*.

BENEVENUTO, F., RODRIGUES, T., CHA, M., AND ALMEIDA, V. 2009b. Characterizing user behavior in online social networks. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*.

BOLL, S. 2007. Multitube–where Web 2.0 and multimedia could meet. *IEEE MultiMedia 14*, 1, 9–13.

BRESLAU, L., CAO, P., FAN, L., PHILLIPS, G., AND SHENKER, S. 1999. Web Caching and Zipf-like Distributions: Evidence and Implications. In Proceedings of *IEEE Infocom*.

BRIN, S. AND PAGE, L. 1998. The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw. ISDN Syst. 30*, 1–7, 107–117.

BRODER, A., KUMAR, R., MAGHOUL, F., RAGHAVAN, P., RAJAGOPALAN, S., STATA, R., TOMKINS, A., AND WIENER, J. 2000a. Graph structure in the Web. *Comput. Netw. 33*, 1–6, 309–320.

BRODER, A., KUMAR, R., MAGHOUL, F., RAGHAVAN, P., RAJAGOPALAN, S., STATA, R., TOMKINS, A., AND WIENER, J. 2000b. Graph structure in the Web: Experiments and models. In *Proceedings of the International World Wide Web Conference (WWW)*.

CHA, M., KWAK, H., RODRIGUEZ, P., AHN, Y., AND MOON, S. 2007. I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*.

CHAU, D., PANDIT, WANG, S., AND FALOUTSOS, C. 2007. Parallel crawling for online social networks. In *Proceedings of the International World Wide Web Conference (WWW)*.

CHOUDHURY, M., SUNDARAM, H., JOHN, A., AND SELIGMANN, D. 2009. What makes conversations interesting? themes, participants and consequences of conversations in online social media. In *Proceedings of the International World Wide Web Conference (WWW)*.

CHUA, T.-S., TANG, J., HONG, R., LI, H., LUO, Z., AND ZHENG, Y.-T. 2009. Nus-wide: A real-world Web image database from National University of Singapore. In *Proceedings of the ACM Conference on Image and Video Retrieval (CIVR'09)*.

CHUN, H., KWAK, H., EOM, Y., AHN, Y., MOON, S., AND JEONG, H. 2008. Comparison of online social relations in volume vs interaction: A case study of cyworld. In *Proceedings of the ACM SIGCOMM Conference on Internet Measurement (IMC)*.

DUARTE, F., MATTOS, B., ALMEIDA, J., ALMEIDA, V., CURIEL, M., AND BESTAVROS, A. 2008. Hierarchical characterization and generation of blogosphere workloads. Tech. rep., Computer Science Department, Boston University.

EBEL, H., MIELSCH, L., AND BORNHOLDT, S. 2002. Scale free topology of e-mail networks. *Phys. Rev. E 035103*.

GARLASCHELLI, D. AND LOFFREDO, M. 2004. Patterns of link reciprocity in directed networks. *Phys. Rev. Lett. 93*.

GILL, P., ARLITT, M., LI, Z., AND MAHANTI, A. 2007. Youtube traffic characterization: A view from the edge. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*.

GOLDER, S., WILKINSON, D., AND HUBERMAN, B. 2007. Rhythms of social interaction: Messaging within a massive online network. *Proceedings of the International Conference on Communities and Technologies*.

GOMES, L., ALMEIDA, J., ALMEIDA, V., AND MEIRA, W. 2007. Workload models of spam and legitimate e-mails. *Perform. Eval. 64*, 7-8, 690–714.

GÓMEZ, V., KALTENBRUNNER, A., AND LÓPEZ, V. 2008. Statistical analysis of the social network and discussion threads in slashdot. In *Proceedings of the International Conference on World Wide Web (WWW)*.

HEYMANN, P., KOUTRIKA, G., AND GARCIA-MOLINA, H. 2007. Fighting spam on social Web sites: A survey of approaches and future challenges. *IEEE Internet Comput. 11*.

HUBERMAN, B., ROMERO, D., AND WU, F. 2009. Social networks that matter: Twitter under the microscope. *First Monday 14*, 1.

JAIN, R. 1991. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. John Wiley and Sons.

JAVA, A., SONG, X., FININ, T., AND TSENG, B. 2007. Why we Twitter: Understanding microblogging usage and communities. In *Proceedings of the Workshop on Web Mining and social network analysis (WebKDD/SNA-KDD)*.

KLEINBERG, J. 1999. Hubs, authorities, and communities. *ACM Comput. Surv. 31*, 5.

KRISHNAMURTHY, B., GILL, P., AND ARLITT, M. 2008. A few chirps about Twitter. In *Proceedings of the ACM SIGCOMM Workshop on Social Networks (WOSN)*.

KUMAR, R., NOVAK, J., RAGHAVAN, P., AND TOMKINS, A. 2003. On the bursty evolution of blogspace. In *Proceedings of the International World Wide Web Conference (WWW)*.

LESKOVEC, J. AND HORVITZ, E. 2008. Planetary-scale views on a large instant-messaging network. In *Proceedings of the International Conference on World Wide Web (WWW)*.

LIBEN-NOWELL, D., NOVAK, J., KUMAR, R., RAGHAVAN, P., AND TOMKINS, A. 2005. Geographic routing in social network. In *Proc. Nat. Aca. Sci.*

MISLOVE, A., MARCON, M., GUMMADI, K., DRUSCHEL, P., AND BHATTACHARJEE, B. 2007. Measurement and analysis of online social networks. In *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*.

NEWMAN, M. 2002. Assortative mixing in networks. *Phys. Rev. Lett. 89*.

NEWMAN, M. AND PARK, J. 2003. Why social networks are different from other types of networks. *Phys. Rev. E 68*.

ONNELA, J., SARAMÄKI, J., HYVÖNEN, J., SZABÓ, G., MENEZES, A., KASKI, K., BARABÁSI, A., AND KERTÉSZ, J. 2007. Analysis of a large-scale weighted network of one-to-one human communication. *New J. Phys. 9*, 6 (June).

PANDURANGAN, G., RAGHAVAN, P., AND UPFAL, E. 2002. Using Pagerank to characterize Web structure. In *Proceedings of the International Conference on Computing and Combinatorics (COCOON)*.

ROSS, K. 2003. Asynchronous voice: A personal account. *Proceedings of the IEEE Multimedia 10*, 2, 70–74.

SESHADRI, M., MACHIRAJU, S., SRIDHARAN, A., BOLOT, J., FALOUTSOS, C., AND LESKOVE, J. 2008. Mobile call graphs: Beyond power-law and lognormal distributions. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*.

SHANNON, M. 2007. Shaking hands, kissing babies, and . . . blogging? *Commun. ACM 50*.

TRIVEDI, K. S. 2002. *Probability and Statistics with Reliability, Queuing and Computer Science Applications*. Prentice-Hall.

ZHANG, J., ACKERMAN, M., AND ADAMIC, L. 2007. Expertise networks in online communities: Structure and algorithms. In *Proceedings of the International World Wide Web Conference (WWW)*.

ZINK, M., SUH, K., GU, Y., AND KUROSE, J. 2008. Watch global, cache local: Youtube network traces at a campus network—measurements and implications. In *Proceedings of the IEEE Multimedia Computing and Networking (MMCN)*.

ZINMAN, A. AND DONATH, J. 2007. Is Britney Spears spam? In *Proceedings of the Conference on Email and Anti-Spam (CEAS)*.

ZLATIC, V., BOZICEVIC, M., STEFANCIC, H., AND DOMAZET, M. 2006. Wikipedias: Collaborative Web-based encyclopedias as complex networks. *Phys. Rev. E 74*, 1.

Received January 2009; revised May 2009; accepted June 2009